RESEARCH ARTICLE

WILEY

A novel algorithm for lung adenocarcinoma based on N6 methyladenosine-related immune long noncoding RNAs as a reliable biomarker for predicting survival outcomes and selecting sensitive anti-tumor therapies

Qiuwen Yan[1] | Bingchuan Hu[1] | Hang Chen[1] | Linwen Zhu[1] | Yao Lyu[2] | Dingding Qian[3] | Guofeng Shao[1]

[1]Department of Cardiothoracic Surgery, The Affiliated Lihuili Hospital, Ningbo University, Ningbo, Zhejiang, People's Republic of China

[2]Department of Pharmacy, The Affiliated Lihuili Hospital, Ningbo University, Ningbo, Zhejiang, People's Republic of China

[3]Department of Cardiology, The Affiliated Lihuili Hospital, Ningbo University, Ningbo, Zhejiang, People's Republic of China

Correspondence
Guofeng Shao, Department of Cardiothoracic Surgery, The Affiliated Lihuili Hospital, Ningbo University, Ningbo, Zhejiang, People's Republic of China.
Email: sgf1958@sina.com

Funding information
Natural Science Foundation of Ningbo, Grant/Award Number: 202003N4269;2021J288; Ningbo Medical Science and Technology Plan Project, Grant/Award Number: 2020Y01; Ningbo Medical and Health Brand Discipline, Grant/Award Number: PPXK2018–01; Ningbo "Technology Innovation 2025" Major Special Project, Grant/ Award Number: 2022Z150; Zhejiang Province Basic Public Welfare Research Program Project, Grant/Award Number: LGF22H010006

Abstract

Background: Lung cancer is a highly heterogeneous malignant tumor with high incidence and mortality. Recently, increasing evidence has demonstrated that N6-methyladenosine (m6A) methylation and the tumor microenvironment (TME) play important roles in the occurrence and development of lung adenocarcinoma (LUAD).

Methods: In this study, we constructed a novel and reliable algorithm based on m6A-related immune lncRNAs (mrilncRNAs), consisting of molecular subtypes and a prognostic signature.

Results: According to the analyses of molecular subtypes, patients in cluster 1 were in a more advanced stage, showed poor prognosis, were sensitive to immunotherapy (anti-programmed cell death 1 Ligand 1 (PD-L1) and anti-lymphocyte activating 3 (LAG-3)), and had a highest tumor mutational burden (TMB), while anti-cytotoxic T-lymphocyte-associated protein 4 (CTLA-4) therapy seemed to be a good choice for patients in cluster 3. Subsequently, the results of the risk assessment model indicated that the low-risk patients exhibited a survival advantage, had an earlier stage, and showed a higher response to common anti-cancer drugs, including chemotherapy (Docetaxel, Paclitaxel), molecular targeted therapy (Erlotinib), and immunotherapy (anti-CTLA-4 therapy), while Gefitinib could be a good choice for patients with high-risk scores.

Conclusion: In conclusion, the constructed algorithm exhibits promising practical prospects, and allows the selection of suitable and sensitive anti-cancer drugs, which could provide theoretical support to predict the survival outcomes of patients with LUAD.

KEYWORDS
consensus cluster, immune, LncRNAs, M6A methylation, prognostic signature

---

1 | INTRODUCTION

Lung cancer is a highly heterogeneous malignant tumor with high incidence and mortality.[1] As the second-leading cause of cancer-related deaths in both men and women, lung cancer is projected by the American Cancer Society to kill 117,910 men and 118,830 women in the United States in 2022.[2] According to the National Central Cancer Registry of China, lung cancer was the most common malignancy in 2015 and the number one cause of cancer-related death in both men and women.[3] The most common classification of lung cancer is nonsmall cell lung cancer (NSCLC), accounting for about 85% of lung cancer cases, of which more than 60% of the pathological types are lung adenocarcinoma (LUAD).[4] In recent years, there have been continuous breakthroughs in standard treatment strategies for lung cancer, such as chemotherapy, radiotherapy, molecular targeted therapy, immunotherapy, and traditional Chinese medicine, with surgery remaining the first choice for early-stage lung cancer treatment.[5] However, the 5-year survival rate for patients with advanced NSCLC with distant metastasis is only 7%.[6] Therefore, faced with such a severe situation, for patients with advanced lung cancer, selecting sensitive anti-tumor drugs and providing targeted treatment are very significant.

As a common reversible RNA internal modification, N6-methyladenosine (m6A) methylation is closely associated with NSCLC initiation and progression.[7] m6A methylation is closely associated with fundamental cellular functions, including 3′-end and microRNA (miRNA) processing, pre-mRNA splicing, and translation regulation.[8] Chen et al. introduced the mechanism of m6A methylation in detail and provided a list of m6A-related genes, which were utilized in this study.[9] Wang et al. recommended that targeting methyltransferase 14, N6-adenosine-methyltransferase subunit (METTL14), an mRNA methylase, could be a promising therapy to hinder the progression of malignancies.[10] Ma et al. proposed that the occurrence and development of LUAD was closely associated with the regulation of m6A methylation.[11] In recent years, research on m6A methylation-related genes has become a research hotspot.[12,13] However, long noncoding RNAs (lncRNAs), comprising noncoding RNAs with transcripts of >200 nucleotides, related to m6A methylation have not received sufficient attention. Therefore, we wonder whether m6A-related immune lncRNAs (mrilncRNAs) are related to the occurrence and development of lung cancer, and whether they can be utilized as biomarkers for predicting the prognosis of LUAD and selecting anti-tumor drugs.

The tumor microenvironment (TME) plays an important regulatory role in the occurrence and progression of NSCLC.[14] Interaction among its internal components allow the TME to influence tumor growth and drug delivery in a highly complex manner.[15] Furthermore, an immunosuppressive TME could substantially reduce treatment response and overall survival (OS) of patients with malignant tumors.[16] To date, studies have reported bioinformatic analyses of immune-related genes and lncRNAs[17,18]; however, such analysis of mrilncRNAs in patients with LUAD is lacking. Therefore, we speculate that mrilncRNAs may be involved in the progression of LUAD by regulating TME in patients with LUAD.

In this study, we constructed a novel and reliable algorithm based on mrilncRNAs combining molecular subtypes and a prognostic signature, which could provide theoretical support to predict the survival outcomes of patients with LUAD and allows the selection of suitable and sensitive anti-cancer drugs. Furthermore, the algorithm classified the patients with LUAD twice, representing a step closer to personalized treatment.

2 | MATERIALS AND METHODS

2.1 | Data acquisition

We downloaded the transcriptome data and clinical data of patients with LUAD from The Cancer Genome Atlas (TCGA, https://portal.gdc.cancer.gov/),[19] and distinguished between mRNAs and lncRNAs using gene transfer format (GTF) files from Ensembl (http://asia.ensembl.org).[20] Then, lists of immune-related genes were obtained from The Immunology Database and Analysis Portal (ImmPort, https://www.immport.org/).[21] Subsequently, immune-related lncRNAs (correlation [|cor|]>0.6 and $p<0.001$) and m6A-related lncRNAs (|cor|>0.4 and $p<0.001$) were identified using Spearman correlation analysis by R-x64-4.0.4 limma package between mRNAs and lncRNAs. Next, the mrilncRNAs were screened by taking the intersection m6A-related and immunity-related lncRNAs, and the results were visualized using a Venn diagram. Furthermore, univariate Cox analysis by R-x64-4.0.4 survival package was conducted to screen the mrilncRNAs related to the survival of patients of LUAD for subsequent analyses.

The 16 mrilncRNAs included in the modeling process were visualized using a network diagram and a forest map, in which the network diagram was generated using Cytoscape version 3.8.2.[22]

2.2 | Cluster analyses of mrilncRNAs

Consensus cluster analyses were conducted to divide the patients with LUAD into different clusters by running the ConsensusClusterPlus package.[23] A survival curve was generated to exhibit the overall survival of patients with LUAD from different molecular subtypes by R-x64-4.0.4 survival, survminer packages. Subsequently, the clinical heatmap not only showed the expression of mrilncRNAs related to prognosis, but also revealed the potential relationships between the molecular subtypes and common clinicopathological characteristics (e.g., N, M, T, stage, sex, and Age) by conducting a series of Wilcoxon rank-sum tests. Wilcoxon rank-sum tests were conducted to explore the mRNA expression levels of common immune checkpoint inhibitors (ICIs) (e.g., programmed cell death 1 Ligand 1 (PD-L1), lymphocyte activating 3 (LAG-3), and cytotoxic T-lymphocyte-associated protein 4 (CTLA-4)) in patients with LUAD between different molecular subtypes by R-x64-4.0.4 limma, ggplot2, ggpubr packages, which were labeled as: ***<0.001, **<0.01, and *<0.05. To explore the TEM of the different molecular subtypes, the StromalScore,

ImmuneScore, and ESTIMATEScore (StromalScore + ImmuneScore) were calculated by running the R-x64-4.0.4 estimate package, the statistically differences of which between different molecular subtypes were explored using Wilcoxon rank-sum tests. Furthermore, the TMB is the total number of mutations in million bases in a tumor. Thus, Wilcoxon rank-sum tests were performed to explore whether the TMB of patients with LUAD between different molecular subtypes was statistically significantly different.

2.3 | Construction of a prognostic signature based on mrilncRNAs

The patients with LUAD were randomized into two groups evenly, the training group and the test group. The data from patients in the training group were subjected to least absolute shrinkage and selection operator (LASSO) regression and multivariate Cox regression analyses (Table 1), successively, and the risk scores of patients with LUAD were calculated based on the following formula:

$$\text{Risk score} = \sum_{i=1}^{16} Coef\,(i) \times E(i),$$

In the formula, Coef (i) and E(i) represent the regression coefficients and the expression of the 16 mrilncRNAs included in multivariate Cox regression, respectively. Based on the median risk scores, patients with LUAD in training group were divided into high- and low-risk groups. According to the formula and the cut-off point derived from the training group, the patients in the test group were classified into different risk groups. Furthermore, a Sankey diagram was generated by R-x64-4.0.4 ggalluvial package as a good representation

TABLE 1 The regression coefficients of mrilnRNAs included in the multivariate Cox regression

ID	Coef
AC026462.3	0.0787201295515719
SALRNA1	−0.0156143768702813
AP001178.1	0.000204311904485358
AC004704.1	0.0185730743535504
AC010999.2	−0.690318746272766
AC073316.3	−0.270359798410301
AC009226.1	0.713638807528353
AL031667.3	0.0669255404198781
AC008957.1	−0.188902832545078
AC007613.1	−0.972775984122169
AC005884.1	−0.151433935466164
LINC02587	0.0151465469498612
AC016747.2	−0.0491510127311617
AC026355.2	−0.0912193157895403
AC018529.1	−0.443311743610806
AL137003.1	−0.082228989521466

of the relationship between the molecular subtypes and the prognostic signature. To validate the predictive capability of the risk assessment model, two 1-year receiver operating characteristic (ROC) curves were plotted by running timeROC package and corresponding areas under the ROC curve (AUC) values were calculated. Then, two survival curves were generated to explore the survival outcomes of the patients with LUAD, and four scatter plots were plotted to further reveal the potential relationship between the risk score and patient survival. To investigate whether the risk assessment model could act as a reliable independent prognostic indicator for patients with LUAD, univariate and multivariate Cox regression analyses were conducted for validation, and four forest maps were generated for visualization. The clinical heatmap and a series of box-plots exhibited the relationship between the risk groups, molecular subtypes, and the common clinicopathological characteristics (e.g., N, M, T, stage, sex, age, and ImmuneScore), which were labeled as: ***<0.001, **<0.01, and *<0.05. Subsequently, Wilcoxon rank-sum tests were employed to explore whether there were statistical differences in the expression levels of common ICIs between the different risk groups. In addition, Immunophenoscores (IPS) representing the efficacy of anti-PD-1/CTLA-4 immunotherapy were downloaded from The Cancer Immunome Atlas (TCIA, https://tcia.at/) database.[24] Next, the median inhibitory concentration (IC50) of common anti-tumor drugs (e.g., Docetaxel, Paclitaxel, Gefitinib, and Erlotinib) were calculated by running the R-x64-4.0.4 pRRophetic package.[25] Thus, ideal anti-tumor drugs could be identified using Wilcoxon rank-sum tests to treat patients from different risk groups. To investigate the survival outcomes of patients with LUAD from different molecular subtypes and risk groups, we generated a survival curve for visualization. Furthermore, to exhibit the 1, 3, and 5-year survival rates of patients intuitively, we plotted a nomogram and two correlation curves using the R-x64-4.0.4 rms package for visualization.

3 | RESULTS

3.1 | Identification of mrilncRNAs

The flowchart for the construction of the novel algorithm is shown in Figure 1. The transcriptome data and corresponding clinical information of 551 tissues (54 normal tissues and 497 LUAD tissues) were downloaded from the TCGA database. A total of 23 m6A methylation-related genes, 2483 immune-related genes, and 14,087 lncRNAs were obtained. Subsequently, 1985 m6A methylation-related lncRNAs, and 1952 immune-related lncRNAs were identified by Spearman correlation analyses. Then, 1619 mrilncRNAs were identified by taking the intersection of the m6A methylation-related lncRNAs and immune-related lncRNAs. Ultimately, 94 mirlncRNAs related to prognosis were filtered by performing univariate Cox analysis (Figure 2A), of which 16 were selected for subsequent multivariate Cox analysis in the modeling process (Figure 2B, C). A heatmap was used to exhibit the expression levels of the 94 mirlncRNAs related to prognosis in normal and LUAD tissues (Figure 2D).
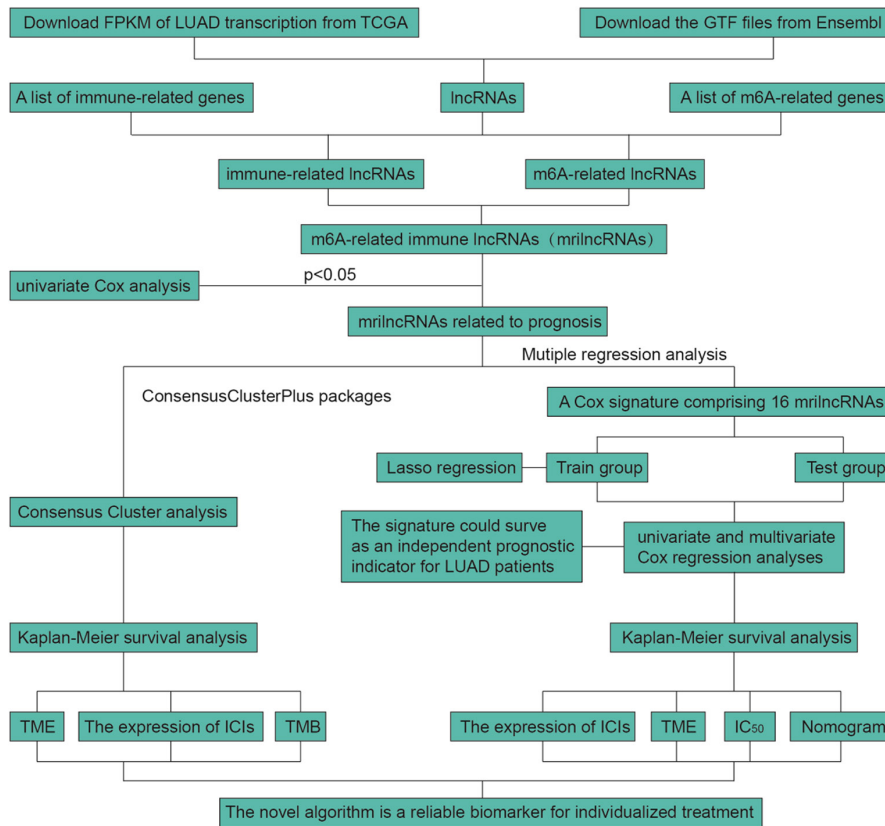
3.2 | The molecular subtypes based on mirlncRNAs are reliable biomarkers for selecting suitable immunotherapy

The 94 mirlncRNAs related to prognosis were subjected to consensus cluster analyses, generating 9 consensus matrices. When the k value was set to 3, the constructed molecular subtypes exhibited the best results (Figure 3A–C). Therefore, patients with LUAD were divided into three molecular subtypes based on the expression levels of the 94 mirlncRNAs. The survival curve suggested that patients in cluster 3 exhibited the best survival advantages, followed by patients in cluster 2, while the patients in cluster 1 had the worst survival (Figure 3D). The clinical heatmap showed the expression levels of the 94 mirlncRNAs of LUAD tissues, and revealed that the molecular subtypes were closely associated with the Stage (Figure 4A), in which more patients with advanced Stage disease were in cluster 1. According to the expression analyses of ICIs, anti-PD-L1 (Figure 4B), and anti-LAG-3 (Figure 4D) therapies could be good treatment strategies for patients in cluster 1, while anti-CTLA 4 therapy (Figure 4C) could be a good choice for patients in cluster 3. According to the TME, the patients with LUAD in cluster 2 tended to have a relative higher StromalScore (Figure 4E), ImmuneScore (Figure 4F), and ESTIMATEScore (Figure 4G), indicating that they had a relative higher abundance of stromal cells and immune cells. The TMB, as a biomarker, can effectively predict patients' response to immunotherapy. The TMB of patients in cluster 1 was significantly higher than that of the other two clusters, suggesting that the patients in cluster 1 could be more sensitive to immunotherapy (Figure 5A). In

conclusion, compared with those in the other two clusters, patients in cluster 1 could be more suitable for immunotherapy, including anti-PD-L1 or anti-LAG-3 therapies.

3.3 | The risk assessment model is a robust tool to select sensitive anti-cancer drugs

The patients in the training group were included in LASSO regression followed by multivariate Cox regression (Figure 5B), and 16 mirlncRNAs were selected to construct the prognostic signature. The Sankey diagram showed that the majority of patients in cluster 1 were high-risk patients, while most of the cluster 3 patients were patients with low-risk scores, and in cluster 2, high-risk and low-risk patients were almost equally divided (Figure 5C). The 1-year AUC values of the training group and test group were 0.800 (Figure 5D), and 0.715 (Figure 5E), respectively. According to the Kaplan–Meier survival analyses of the two groups, we concluded that high-risk patients tended to have relatively worse prognosis (Figure 5F, G). Furthermore, there were significantly more deaths among high-risk patients than among low-risk patients (Figure 6A, B). According to univariate and multivariate Cox regression analyses, we discovered that two factors, Stage and risk score, could act as independent prognostic indicators in patients with LUAD (Figure 6C, D). The clinical heatmap (Figure 7A) and a series of box-plots suggested that the risk score was closely related to N (Figure 7B), T (Figure 7C), Stage (Figure 7D), ImmuneScore (Figure 7E), and Cluster (Figure 7F), in which low-risk patients tended to have an earlier stage of N, T, and
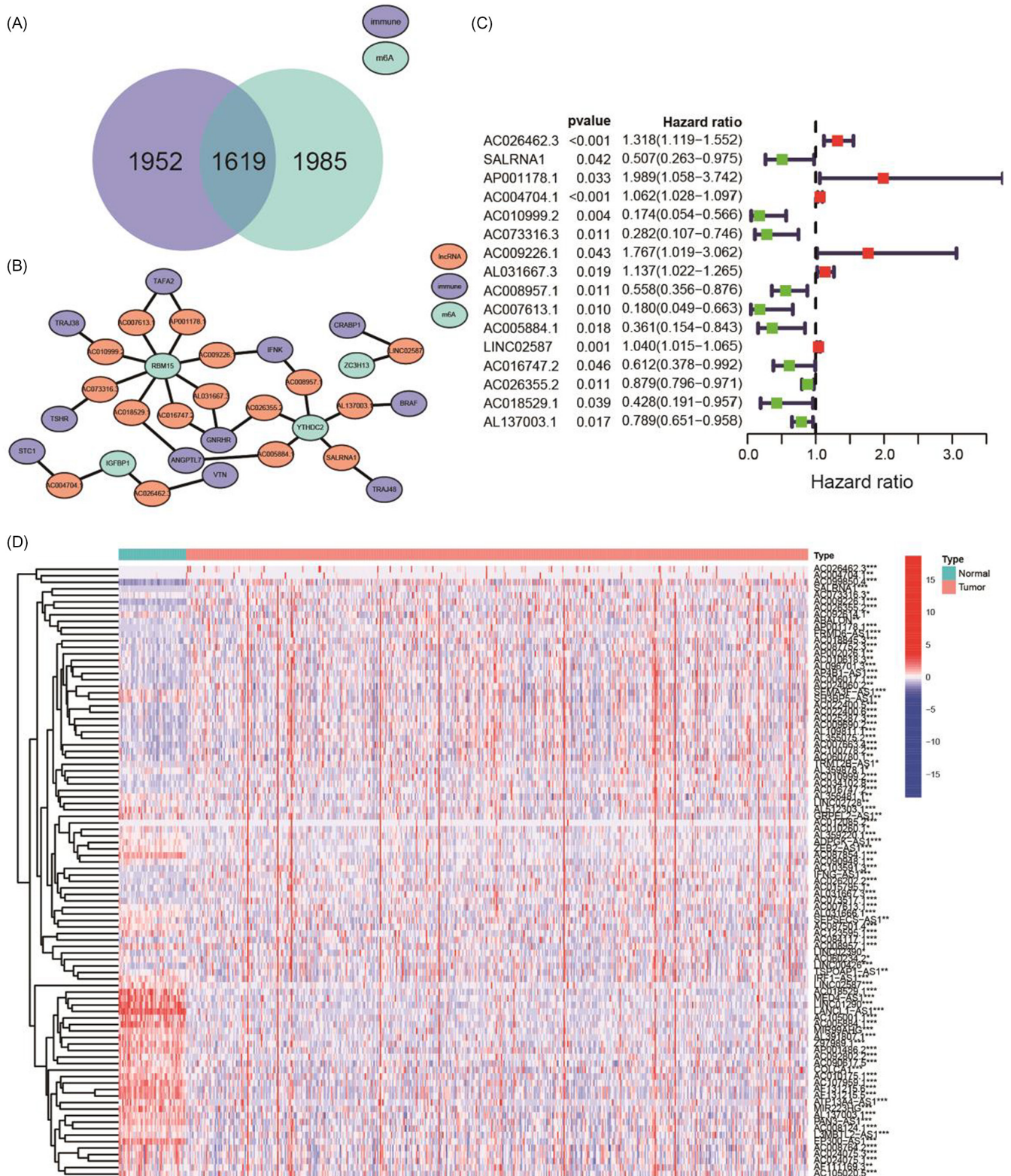
FIGURE 2 Identification of the mrilncRNAs. (A) 1985 m6A methylation-related lncRNAs, and 1952 immunity-related lncRNAs were identified by Spearman correlation analyses; 1619 mrilncRNAs were subsequently identified by taking the intersection the two sets of lncRNAs. (B) Network diagram revealing the relationship between the m6A-related lncRNAs, immune-related lncRNAs, and 16 mirlncRNAs (only the most relevant lncRNAs are shown). (C) Forest map showing that the 16 mrilncRNAs are closely associated with survival, and most of them are protective genes. (D) Heatmap showing the expression levels of 94 mrilncRNAs related to prognosis in normal and LUAD tissues
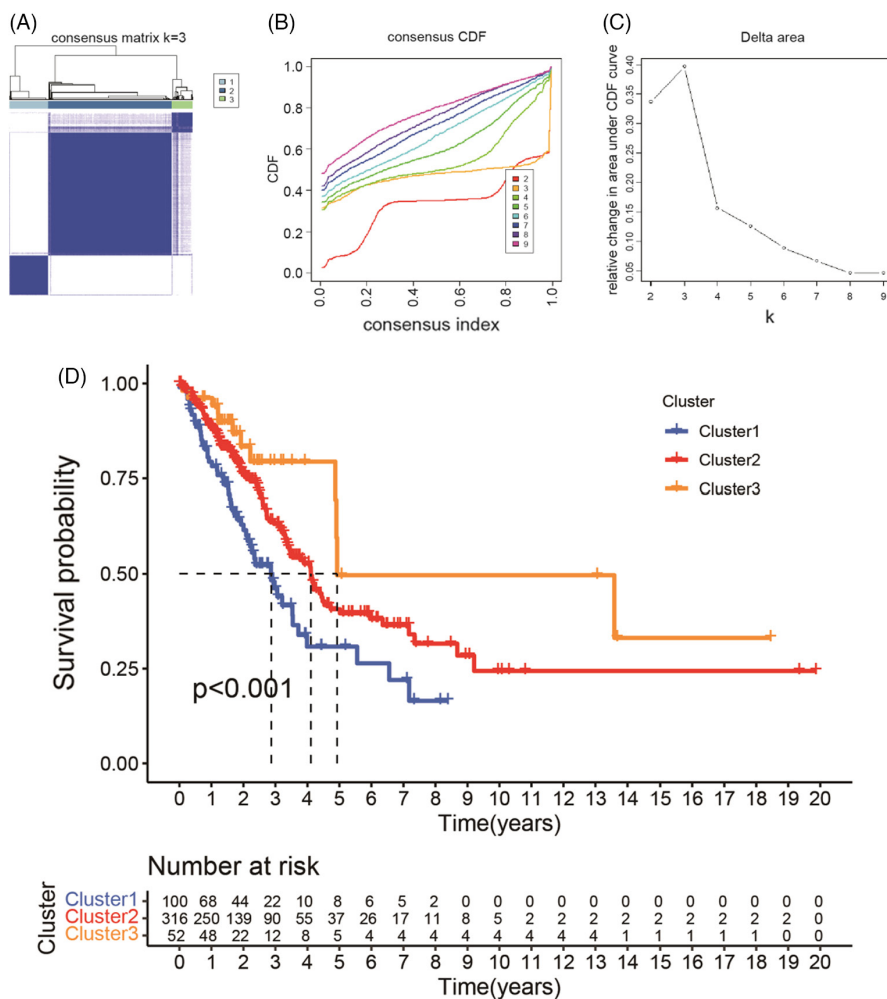
FIGURE 3 Cluster analysis based on 94 mirlncRNAs related to prognosis: Survival analysis. When the k value was set to 3, the ideal molecular subtypes were obtained (A), the slope of the cation diffusion facilitator (CDF) was the minimum (B), and the relative change in the area under the CDF curve was the maximum (C). (D) Survival curve suggesting that patients in cluster 3 exhibited the best survival advantages, followed by patients in cluster 2, while the patients in cluster 1 had the worst survival

Stage, a higher ImmuneScore, and comprised the majority of cluster 2 or cluster 3. Expression analysis associated with IPS analyses indicated that compared with patients with high-risk scores, low-risk patients had a relative higher expression of CTLA4 (Figure 8A) and higher IPS scores for anti-CTLA-4 (Figure 8C), while there was no statistical difference in the IPS of anti-PD-1 between the two groups (Figure 8B), indicating that low-risk patients were more sensitive to anti-CTLA-4 therapy. Moreover, low-risk patients had a relatively higher StromalScore (Figure 8D), ImmuneScore (Figure 8E), and ESTIMATEScore (Figure 8F), suggesting they had a higher content of stromal cells and immune cells. According to the IC50 predicted by the pRRophetic package, we found that the low-risk patients tended to be sensitive to Docetaxel (Figure 8G), Paclitaxel (Figure 8H), and Erlotinib (Figure 8J), while patients with high-risk scores always exhibited a higher response to Gefitinib (Figure 8I). The survival curve revealed that the OS of patients in the different groups were statistically significantly different, and high-risk patients in cluster 1 exhibited the worst survival outcomes (Figure 8K). The nomogram could rapidly predict the 1, 3, and 5-year survival rates of patients with LUAD based on a series of common clinicopathological characteristics (Figure 9A), displaying an excellent predictive capability for 1-year (Figure 9B) and 3-year survival rates (Figure 9C).

4 | DISCUSSION

In recent years, with the development of bioinformatic analyses, research has established reliable prognostic signatures based on mRNAs or lncRNAs to predict the prognosis of patients with various malignancies.[26,27] We were eager to build a new algorithm to predict the survival outcomes of patients with LUAD and provide valuable recommendations for the selection of appropriate anti-tumor drugs.

In this study, we constructed a novel algorithm based on mrilncRNAs, which comprises molecular subtypes and a risk assessment model. We divided patients with LUAD into three clusters by consensus cluster analysis based on the expression of 94 mirlncRNAs, which exhibited the best effects. According to the clinical heatmap, the molecular subtypes were closely associated with the Stage, in which there were more advanced stage patients in cluster 1. Furthermore, the patients in cluster 3 exhibited the best survival advantages, followed by patients in cluster 2, while the patients in cluster 1 had the worst. The expression analyses of ICIs indicated that anti PD-L1 and anti-LAG-3 therapies were suitable for patients in cluster 1, while patients in cluster 3 showed higher responses to anti-CTLA-4 therapy. Moreover, the TMB of patients in cluster 1 was significantly higher than the other two clusters. In summary, the patients with cluster 1 had the most advanced Stage, leading to the worst prognosis. In addition, we found that they had a higher response
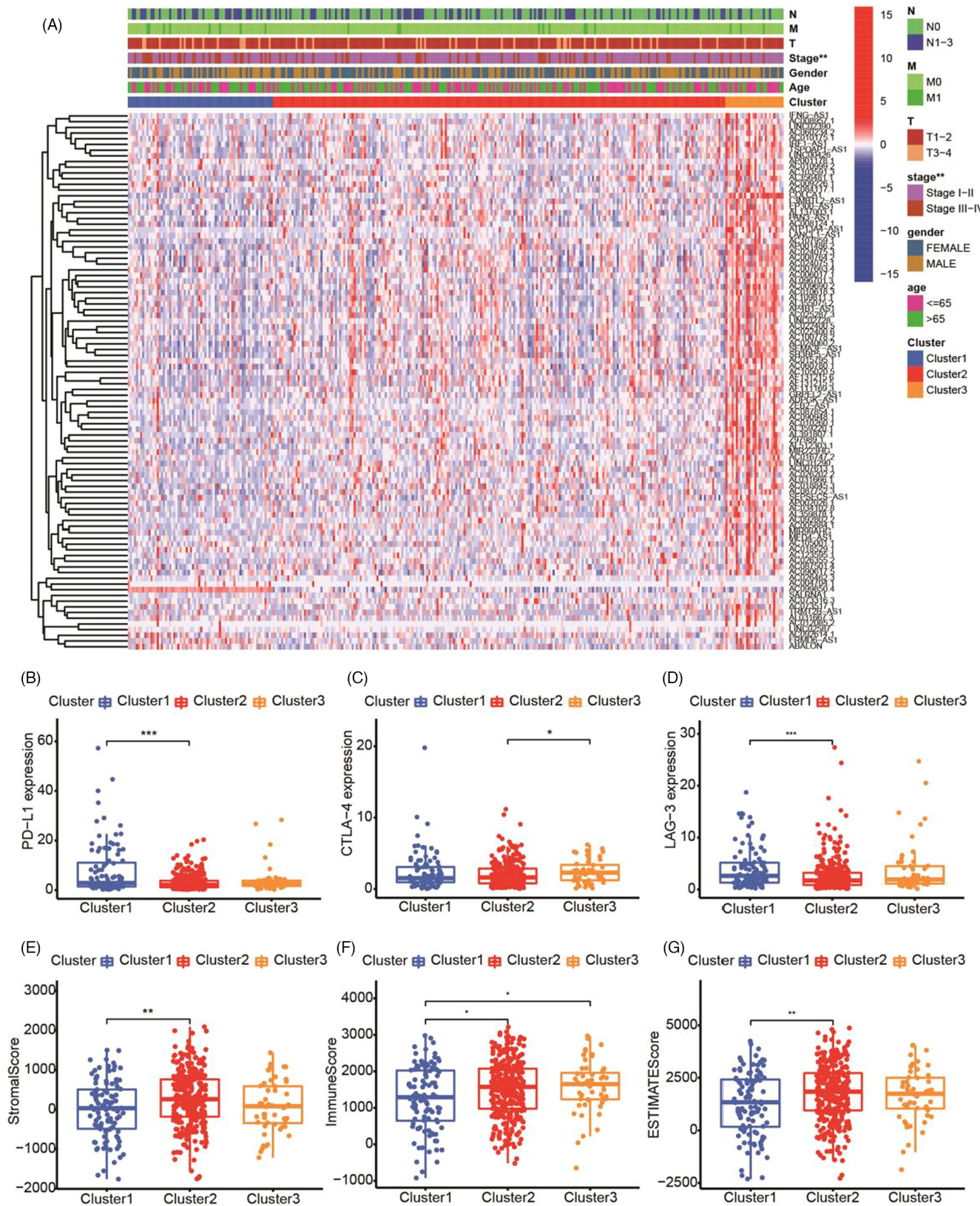
FIGURE 4 Cluster analysis based on 94 mirlncRNAs related to prognosis: Stage and molecular subtypes. (A) The molecular subtypes are closely associated with the Stage, in which there are more advanced Stage patients in cluster 1. The expression levels of PDL1 (B) and LAG3 (D) in patients from cluster 1 were relative higher, while the expression of CTLA4 (C) in patients from cluster 3 was relative higher. The patients with LUAD in cluster 2 always had a relative higher StromalScore (E), ImmuneScore (F), and ESTIMATEScore (G)
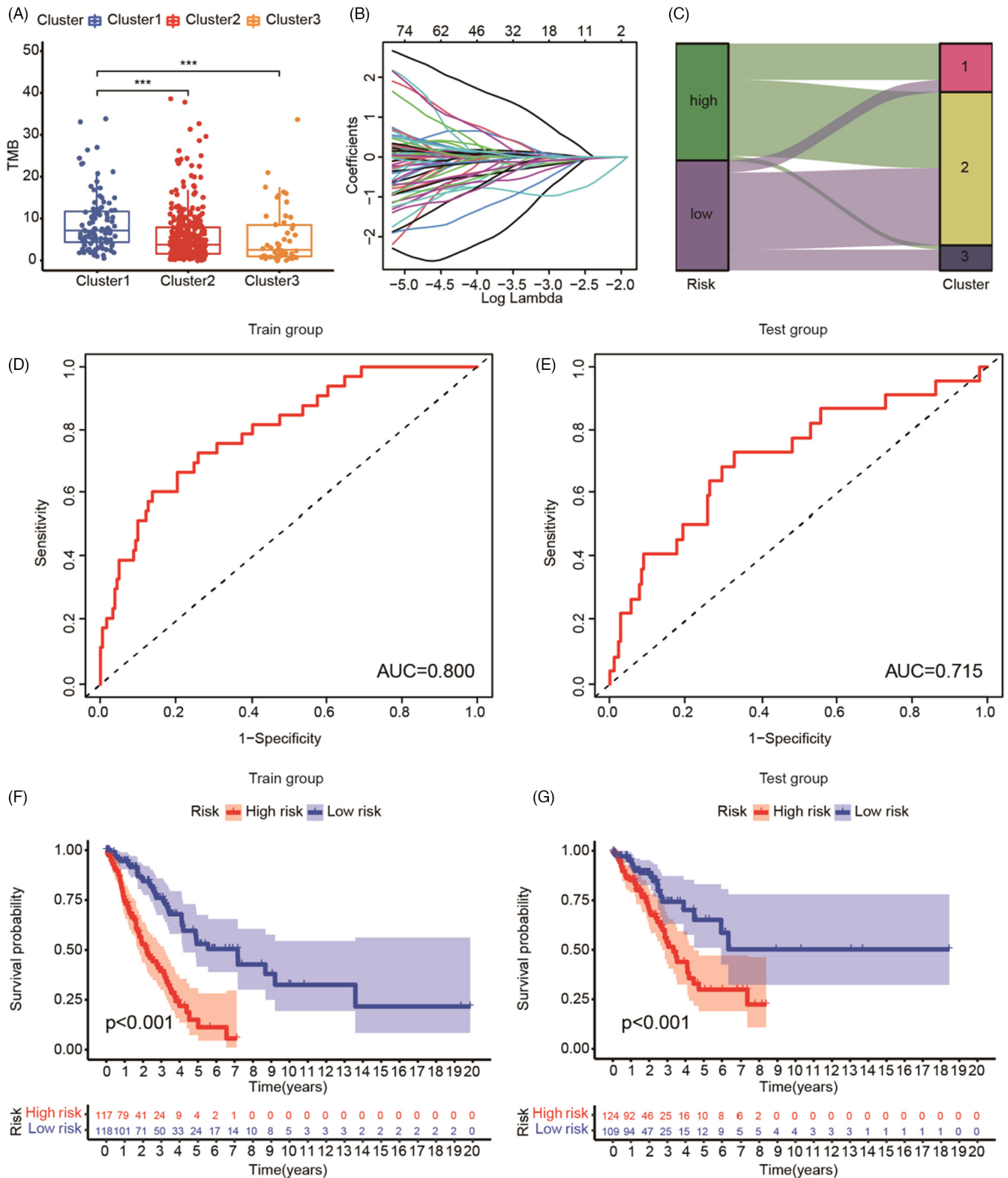
FIGURE 5 Construction of the risk assessment model. (A) The TMB of patients in cluster 1 was significantly higher than that in the other two clusters. (B) The patients in the training group were included in lasso regression followed by multivariate Cox regression analysis. (C) Sankey diagram showing that the majority of patients in cluster 1 are high-risk patients, while most of cluster 3 patients are patients with low-risk scores, and in cluster 2, the high-risk and low-risk patients are almost equally divided. The 1-year AUC values of the training group and the test group were 0.800 (D), and 0.715 (E), respectively. High-risk patients in the training group (F) and test group (G) tend to have a relatively worse prognosis
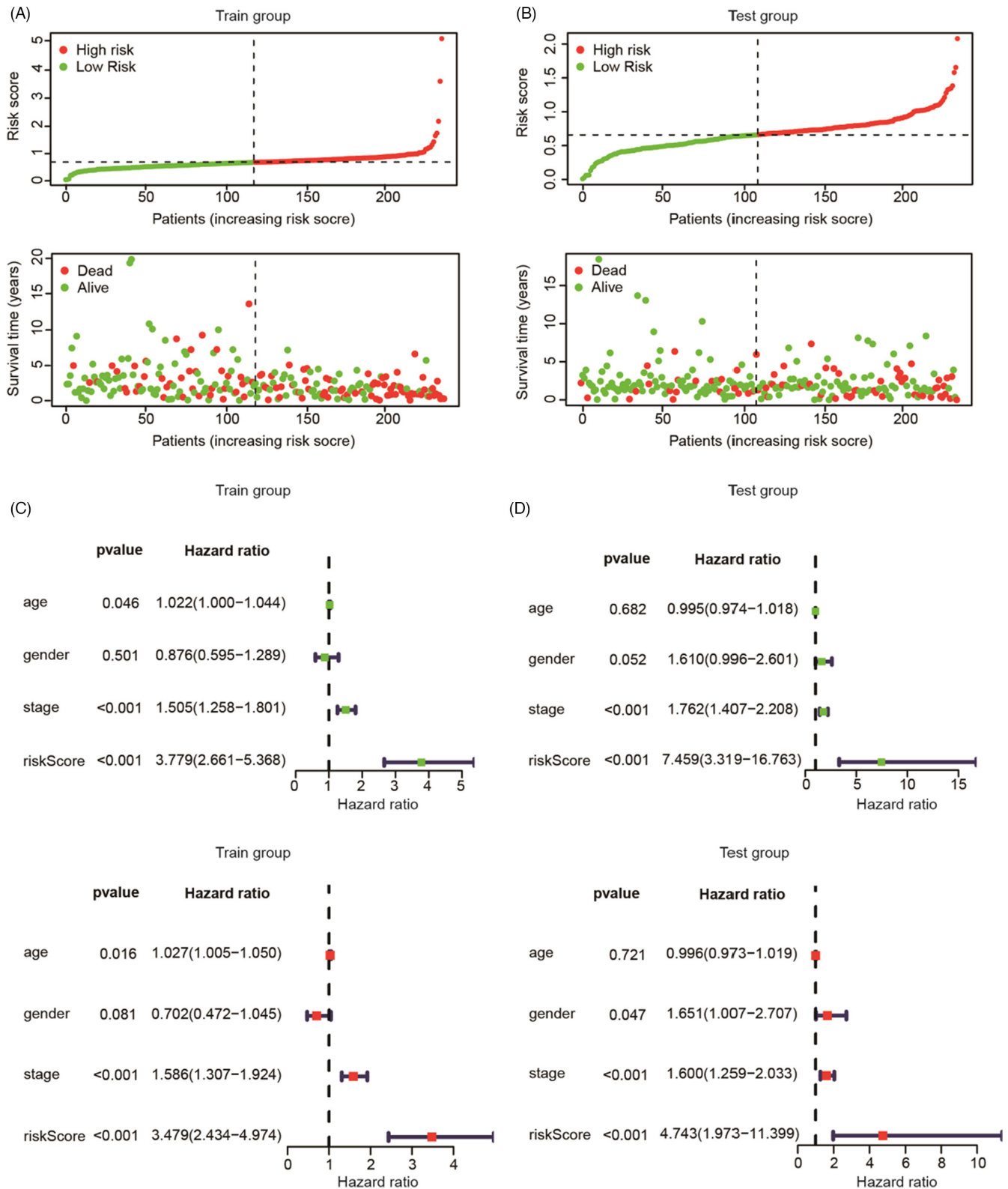
FIGURE 6 Assessment of the predictive capability of the risk assessment model. Significantly more high-risk patients in the training group (A) and the test group (B) died compared with the low-risk patients. (C, D) According to univariate and multivariate Cox regression analyses, two factors, Stage and risk score, could act as independent prognostic indicators in patients with LUAD

to anti-PD-L1 and anti-LAG-3 therapies through expression analyses of ICIs, which was confirmed by analysis of the TMB, indicating that they had a better efficacy of immunotherapy. In conclusion, patients in

cluster 1 showed poor prognosis, and were sensitive to immunotherapy (anti-PD-L1 and anti-LAG-3 therapies), while anti-CTLA-4 therapy seemed to be a good choice for patients in cluster 3.
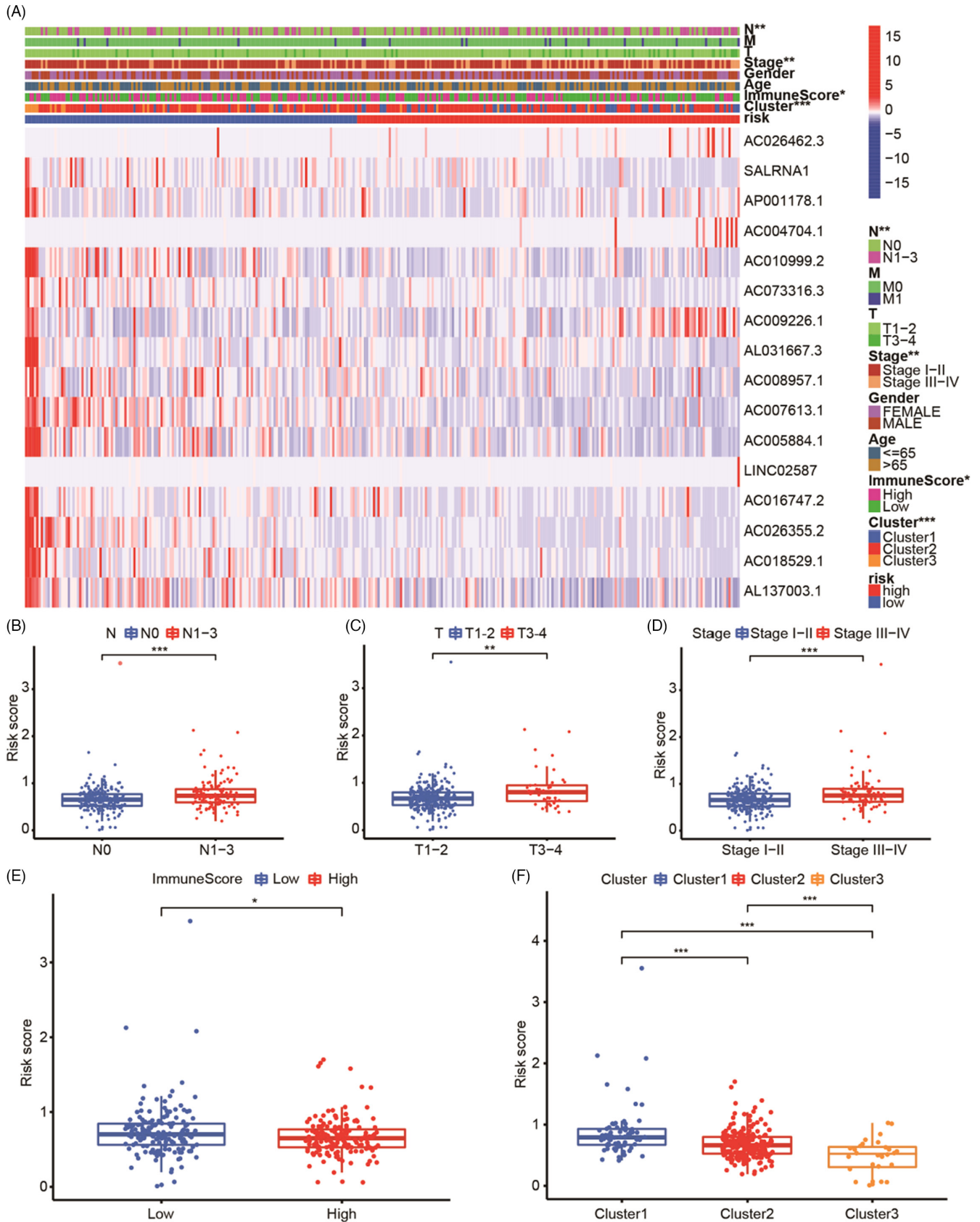
FIGURE 7 Exploration of the clinical significance of the risk assessment model: cancer stage, immune score, and clustering. (A–F) The clinical heatmap (A) and a series of box-plots suggested that the risk score was closely related to N (B), T (C), Stage (D), ImmuneScore (E), and Cluster (F), in which low-risk patients tended to have an earlier stage of N, T, and Stage, a higher ImmuneScore, and comprised the majority of cluster 2 or cluster 3
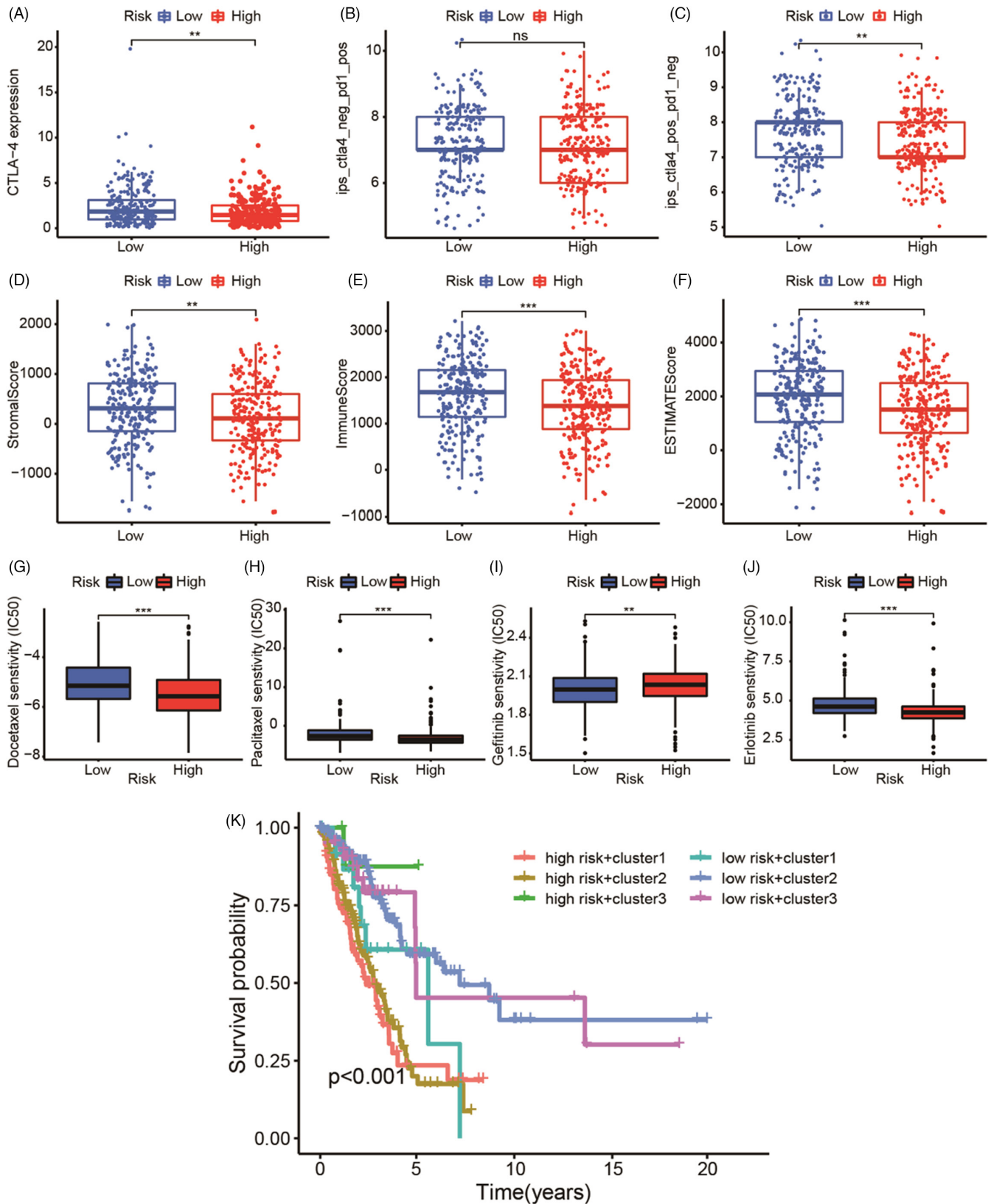
FIGURE 8 Exploration of the clinical significance of the risk assessment model: Immune checkpoint gene expression, stromal and immune scores, and response to chemotherapy. (A–C) Low-risk patients have relative higher expression of CTLA4 (A) higher IPS scores of anti-CTLA-4 (C), whereas there was no statistical difference in IPS of anti-PD-1 between the two groups (B). (D–F) Low-risk patients have a relatively higher StromalScore (D), ImmuneScore (E), and ESTIMATEScore (F). Low-risk patients tend to be sensitive to Docetaxel (G), Paclitaxel (H), and Erlotinib (J), while patients with high-risk scores exhibit a higher response to Gefitinib (I). The OS of patients in the different groups were statistically significantly different, and high-risk patients in cluster 1 exhibited the worst survival outcomes (K)
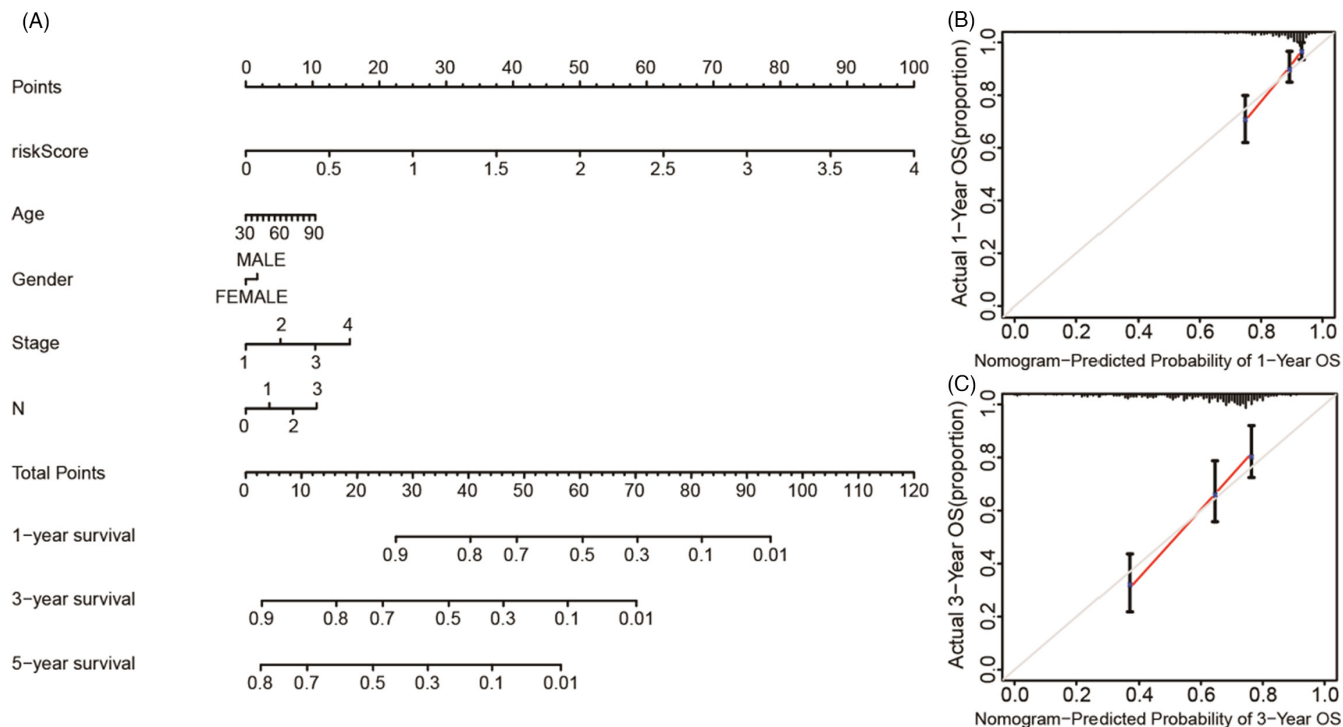
FIGURE 9 Exploration of the clinical significance of the risk assessment model: Survival prediction. (A–C) The nomogram can rapidly predict the 1, 3, and 5-year survival rates of patients with LUAD based on a series of common clinicopathological characteristics (A), which displays an excellent predictive capability for 1-year (B), and 3-year survival rates (C)

Then, patients with LUAD were divided into high- and low-risk groups based on 16 mrilncRNAs after being divided randomly into the training group and the test group randomly. We found that patients in cluster 1 had the highest risk score, followed by patients in cluster 2, while the patients in cluster 3 had the lowest. Moreover, low-risk patients tended to have an earlier stage of N, T, and Stage, a higher StromalScore and ImmuneScore, and a relatively better survival outcomes. We discovered that the low-risk patients had a relatively higher expression of CTLA4, and higher IPS scores of anti CTLA-4, which indicated that they were sensitive to anti-CTLA-4 therapy. Finally, we predicted the IC50 of common drugs, revealing that low-risk patients tended to be sensitive to Docetaxel, Paclitaxel, and Erlotinib, while patients with high-risk scores exhibited a stronger response to Gefitinib. In conclusion, the low-risk patients exhibited a survival advantage, and showed higher responses to common anti-drugs, including chemotherapy (Docetaxel, Paclitaxel), molecular targeted therapy (Erlotinib), and immunotherapy (anti-CTLA-4 therapy), while Gefitinib could be a good choice for patients with high-risk scores.

Among the 16 mrilncRNAs included in the process of modeling, only a few have been reported to be involved in the occurrence and development of multiple malignant tumors. For instance, AC026462.3,[28] AL031667.3,[29] and AC026355.2[19,30] were included in the construction of prognostic signatures for various types of malignancies. However, the remaining lncRNAs: SALRNA1, AP001178.1, AC004704.1, AC010999.2, AC073316.3, AC009226.1, AC008957.1, AC007613.1, AC005884.1, LINC02587, AC016747.2, AC018529.1, and AL137003.1, have never been reported to have

potential relationship with malignancies, and thus might provide reliable evidence for subsequent oncology studies.

In this study, we reported a novel algorithm based on mrilncRNAs, which could provide valuable recommendations to predict survival outcomes and select suitable and sensitive drugs, including chemotherapy (Docetaxel, Paclitaxel), targeted therapy (Gefitinib, Erlotinib), and immunotherapy (PD-L1, CTLA-4, and LAG-3 therapy). Furthermore, compared with traditional single risk assessment models or consensus cluster analysis dividing patients into two risk groups or multiple clusters, the constructed algorithm based on mrilncRNAs combined both of the above analyses, which enabled us to classify patients with LUAD in more detail, and provided more accurate results. In addition, we divided the patients with LUAD into a training group and test group in the process of modeling, which provided internal validation of the reliability of the prognostic signature.

However, there are limitations associated with our study. First, given that the LUAD samples were all downloaded from the same public database, we could not eliminate bias from the profile analyzed. Second, the algorithm ultimately needs to serve doctors clinically. However, the results presented are limited to bioinformatic analyses, and further fundamental experiments, including quantitative real-time reverse transcription-polymerase chain reaction (qRT-PCR) and microarrays, are required to verify its feasibility and reliability.

In summary, the constructed algorithm exhibited promising practical prospects, easily predicting survival outcomes and selecting

sensitive anti-cancer drugs for patients with LUAD merely by detecting the expression of the 16 mrilncRNAs. The algorithm provides theoretical support for doctors to select appropriate treatment strategies, which might represent a step toward personalized treatment for LUAD.

AUTHOR CONTRIBUTIONS

QY and GS contributed to the conception of the study. BH and HC performed the R language analyses. LZ contributed significantly to the analysis and manuscript preparation. QT performed the data analyses and wrote the manuscript. YL and DQ helped to perform the analysis and contributed constructive discussions.

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest.

DATA AVAILABILITY STATEMENT

All data analysed during the current study are accessible from the TCGA database (https://portal.gdc.cancer.gov/).

CONSENT FOR PUBLICATION

Written informed consent for publication was obtained from all participants.

PATIENT CONSENT STATEMENT

Not applicable.

PERMISSION TO REPRODUCE MATERIAL FROM OTHER SOURCES

Not applicable.

CLINICAL TRIAL REGISTRATION

Not applicable.

ORCID

Qiuwen Yan https://orcid.org/0000-0003-3231-0225
Guofeng Shao https://orcid.org/0000-0001-5609-7196

REFERENCES

1. Nasim F, Sabath B, Eapen G. Lung cancer. *Med Clin North Am.* 2019;103(3):463-473.
2. Siegel R, Miller K, Fuchs H, Jemal A. Cancer statistics, 2022. *CA Cancer J Clin.* 2022;72(1):7-33.
3. Chen W, Zheng R, Baade P, et al. Cancer statistics in China, 2015. *CA Cancer J Clin.* 2016;66(2):115-132.
4. Shi J, Wang L, Wu N, et al. Clinical characteristics and medical service utilization of lung cancer in China, 2005–2014: overall design and results from a multicenter retrospective epidemiologic survey. *Lung Cancer.* 2019;128:91-100.
5. Li Z, Feiyue Z, Gaofeng L. Traditional Chinese medicine and lung cancer--from theory to practice. *Biomed Pharmacother.* 2021;137:111381.
6. American Cancer Society Non-small cell lung cancer survival rates by stage [OL]wwwcancerorg/cancer/non-small-cell-lung-cancer/detection-diagnosis-staging/survival-rateshtmlDate last updated: 04/20/21.
7. Dong B, Wu C, Li S, et al. Correlation of m6A methylation with immune infiltrates and poor prognosis in non-small cell lung cancer via a comprehensive analysis of RNA expression profiles. *Annals of Translational Medicine.* 2021;9(18):1465.
8. Zhang L, Luo Y, Cheng T, et al. Development and validation of a prognostic N6-Methyladenosine-related immune gene signature for lung adenocarcinoma. *Pharmacogenomics & Personalized Medicine.* 2021;14:1549-1563.
9. Chen X, Zhang J, Zhu J. The role of mA RNA methylation in human cancer. *Mol Cancer.* 2019;18(1):103.
10. Wang M, Liu J, Zhao Y, et al. Upregulation of METTL14 mediates the elevation of PERP mRNA N adenosine methylation promoting the growth and metastasis of pancreatic cancer. *Mol Cancer.* 2020;19(1):130.
11. Ma L, Xue X, Zhang X, et al. The essential roles of mA RNA modification to stimulate ENO1-dependent glycolysis and tumorigenesis in lung adenocarcinoma. *J Exp Clin Cancer Res.* 2022;41(1):36.
12. Jin D, Guo J, Wu Y, et al. mA demethylase ALKBH5 inhibits tumor growth and metastasis by reducing YTHDFs-mediated YAP expression and inhibiting miR-107/LATS2-mediated YAP activity in NSCLC. *Mol Cancer.* 2020;19(1):40.
13. Jin D, Guo J, Wu Y, et al. mA mRNA methylation initiated by METTL3 directly promotes YAP translation and increases YAP activity by regulating the MALAT1-miR-1914-3p-YAP axis to induce NSCLC drug resistance and metastasis. *J Hematol Oncol.* 2019;12(1):135.
14. Guo Q, Xiao X, Wu C, et al. Clinical roles of risk model based on differentially expressed genes in mesenchymal stem cells in prognosis and immunity of non-small cell lung cancer. *Front Genet.* 2022;13:823075.
15. Kim D, Hwang K, Seo E, et al. Vascularized lung cancer model for evaluating the promoted transport of anticancer drugs and immune cells in an engineered tumor microenvironment. *Adv Healthc Mater.* 2022;11:e2102581.
16. Lin J, Lin J, Weng Y, et al. Radiographical evaluation of tumor immunosuppressive microenvironment and treatment outcomes in gastric cancer: a retrospective, multicohort study. *Ann Surg Oncol.* 2022;29:5022-5033.
17. Chen H, Shen W, Ni S, et al. Construction of an immune-related lncRNA signature as a novel prognosis biomarker for LUAD. *Aging-US.* 2021;13(16):20684-20697.
18. He C, Yin H, Zheng J, Tang J, Fu Y, Zhao X. Identification of immune-associated lncRNAs as a prognostic marker for lung adenocarcinoma. *Transl Cancer Res.* 2021;10(2):998-1012.
19. Tomczak K, Czerwińska P, Wiznerowicz M. The cancer genome atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol.* 2015;19:A68-A77.
20. Howe K, Achuthan P, Allen J, et al. Ensembl 2021. *Nucleic Acids Res.* 2021;49(D1):D884-D891.
21. Bhattacharya S, Andorf S, Gomes L, et al. ImmPort: disseminating data to the public for the future of immunology. *J Immunol Res.* 2014;58:234-239.

22. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 2003;13(11):2498-2504.

23. Wilkerson M, Hayes D. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics*, 2010; 15;26(12):1572-3, 1572, 1573.

24. Prior F, Clark K, Commean P, et al. TCIA: an information resource to enable open science. *Annual International Conference of the IEEE Engineering Medicine and Biology Society.* 2013;2013:1282-1285.

25. Geeleher P, Cox N, Huang R. pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PLoS One.* 2014;9(9):e107468.

26. Chen H, Hu Z, Sang M, et al. Identification of an autophagy-related lncRNA prognostic signature and related tumor immunity research in lung adenocarcinoma. *Front Genet.* 2021;12:767694.

27. Ye W, Huang T. Correlation analysis of m6A-modified regulators with immune microenvironment infiltrating cells in lung adenocarcinoma. *PLoS One.* 2022;17(2):e0264384.

28. Chen H, Pan Y, Jin X, Chen G. Identification of a four hypoxia-associated long non-coding RNA signature and establishment of a nomogram predicting prognosis of clear cell renal cell carcinoma. *Front Oncol.* 2021;11:713346.

29. Zheng Z, Zhang Q, Wu W, et al. Identification and validation of a ferroptosis-related long non-coding RNA signature for predicting the outcome of lung adenocarcinoma. *Front Genet.* 2021;12:690509.

30. Gong Z, Li Q, Li J, Xie J, Wang W. A novel signature based on autophagy-related lncRNA for prognostic prediction and candidate drugs for lung adenocarcinoma. *Transl Cancer Res.* 2022;11(1):14-28.